Group Analysis: Hands-On

Gang Chen

SSCC/NIMH/NIH/HHS

Preview: choosing programs

- Program list
  - 3dttest++, 3dMEMA, 3dANOVAx, 3dMVM, 3dLME
  - 3ttest, 3dRegAna, GroupAna retired
  - Voxel-wise approach
  - ROI analysis not discussed: R, Matlab, Excel, SAS, SPSS
  - uber_ttest.py: for 3ttest++ and 3dMEMA only
  - Other programs: scripting (too hard? $ ---> Rick Reynolds!)
  - Typical mistakes
    - Extra spaces after the continuation character BACKSLASHES (\)
    - Typos
    - Model specifications, misuses of options, …

Preview: choosing programs

- Data structure should not always be the only focus
  - Experiment design: number of explanatory variables (factors and quantitative variables), levels of a categorical variable
  - Balance: equal number of subjects across groups?
  - Missing data: throw out those subjects, or keep the partial data?
  - List all the tests you would like to get out of the group analysis
- If computation cost is of concern
  - Super fast programs: 3dttest++, 3dANOVAx, 3dttest, 3dRegAna
  - Super slow programs: 3dMEMA, 3dMVM, 3dLME, GroupAna
- Special features of 3dMEMA
  - Weights subjects based on reliability
  - Models and identifies outliers at voxel level
  - Handles missing data at voxel level (*e.g.* ECoG data)
  - Cross-subjects variability measures (τ^2, H, I^2, ICC) and group comparisons in τ^2

Preview: learning by 6 examples

- BOLD responses estimated with one basis function
  - 3 groups, 1 numeric variable (between-subjects)
  - ANOVA
  - ANCOVA
  - Within-subject covariate
- BOLD responses estimated with multiple basis functions
  - 1 group
  - 2 groups

Case 1: three groups

- Data information
  - COMT (catechol-*O*-methyl transferase) gene with a Val/Met (valine-to-methionine) polymorphism for schizophrenia
  - 3 genotypic groups: Val/Val (12), Val/Met (10), Met/Met (9)
  - 1 effect estimate from each subject
- What program?
  - Almost everybody immediately jumps to this question!
- Tests of interest?
  - Individual group effects: A, B, and C
  - Pairwise group comparisons: A-B, A-C, and B-C: Two-sample *t*-test
  - Any difference across all three groups? Omnibus *F*-test
- What program?
  - One- or two-sample t-test: 3dttest++, 3dMEMA
  - One-way between-subjects ANOVA: 3dANOVA

Case 1: three groups

- One-way between-subjects ANOVA
  - Each subject has only one response value!
  - GLM, not really a random-effects model:

$$\hat{\beta}_{i(j)} = \alpha_0 + \alpha_1 * x_{1i(j)} + \alpha_2 * x_{2i(j)} + \epsilon_{i(j)}$$

  - Coding for subject grouping: with one group (A) as base (reference) for dummy coding (0s and 1s), α_0 = A, α_1 = B – A, and α_1 = C – A.
  - 3dANOVA
    - Don't directly solve GLM
    - Compute sums of squares: computationally efficient!
  - Alternatives: 3dttest++, 3dMEMA

Case 2: multi-way ANOVA

- Data information
  - 1 subject-grouping variable (Group): young (15) and older (14)
  - 3 within-subject factors:
    - task - 2 levels: Perception and Production
    - Syllable - 2 levels: Simple and Complex
    - Sequence - 2 levels: Simple and Complex
- Tests of interest?
  - Comparisons under various combinations
  - Interactions among the 4 factors
- What program?
  - 3dttest++, 3dMEMA, 3dMVM

Case 3: ANCOVA

- Data information
  - 2 subject-grouping variables
    - Group (2 levels): control () and ssd ()
    - Gender (2 levels): males () and females ()
  - 1 within-subject variable: Condition (4 levels: visWord, visPSW, visCStr, audWord, audPSW)
  - 1 quantitative (between-subjects) variable: Age (mean age not significantly different across groups)
- Tests of interest?
  - Main effects, interactions, various contrasts
- Model $\hat{\beta}_{ij} = \alpha_1 * x_{1j} + ... + \alpha_k * x_{kj} + \delta_i + \epsilon_{ij}$
- What program? 3dMVM, 3dLME

Case 4: Within-subject covariate

- Data information
  - 1 within-subject variable: Condition (2 levels: house, face)
  - 1 quantitative (within-subjects) variable: RT (mean RT not significantly different across conditions)
- Tests of interest?
  - Main effects, interactions, various contrasts
- Model
- What program? 3dLME

$$\hat{\beta}_{ij} = \alpha_1 * x_{1j} + ... + \alpha_k * x_{kj} + \delta_i + \epsilon_{ij}$$

Case 5: one group with multiple basis functions

- Data information
  - 15 subjects
  - One effect of interest modeled with 8 basis (TENT) functions
- Tests of interest?
  - Any overall response at a voxel (brain region)?
- Model $\hat{\beta}_{ij} = \alpha_1 * x_{1j} + ... + \alpha_k * x_{kj} + \delta_i + \epsilon_{ij}$
  - No intercept $\alpha_1 = ... = \alpha_k = 0$
  - Test of interest:
  - Residuals $\boldsymbol{\varepsilon}_{ij}$ are most likely serially correlated
- What program? 3dLME

Case 6: two groups with multiple basis functions

- Data information
  - 15 subjects
  - One effect of interest modeled with 8 basis (TENT) functions
- Tests of interest?
  - Any overall response at a voxel (brain region)?
- Model $\hat{\beta}_{ij} = \alpha_1 * x_{1j} + ... + \alpha_k * x_{kj} + \delta_i + \epsilon_{ij}$
  - No intercept
  - Test of interest: $\alpha_1 = ... = \alpha_k = 0$
  - Residuals $\boldsymbol{\varepsilon}_{ij}$ are most likely serially correlated
- What program? 3dANOVA3 –type 5, 3dMVM (VeenaNair on morgoth?)

Overview: learning by 11 examples

- BOLD responses estimated with one basis function
  - 3 groups
  - 2 conditions
  - 2 conditions with missing data
  - 3 groups + 2 genders
  - 3 groups + 2 conditions
  - 3 groups + 2 genders + 1 numeric variable (between-subjects)
  - 3 groups + 2 conditions + 1 numeric variable (between-subjects)
  - 3 groups + 2 conditions + 2 numeric variables (1 within-subject and 1 between-subjects)
- BOLD responses estimated with multiple basis functions
  - 1 group
  - 2 groups
  - 2 groups + 2 conditions